Stochastic modeling and simulation of biochemical reaction kinetics

by

Animesh Agarwal, M.Tech., B.Tech.

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2011

The Thesis committee for Animesh Agarwal

Certifies that this is the approved version of the following thesis :

Stochastic modeling and simulation of biochemical reaction kinetics

APPROVED BY

SUPERVISING COMMITTEE:

Harel Shouval, Supervisor

Yin Liu

Dedication

This work is dedicated to the loving memory of my grandfather,

Late Shri Revati Raman Agarwal.

His strong belief in my abilities has been a motivation and a guiding force throughout the completion of this work.

Acknowledgments

I wish to thank and express my sincere gratitude to my advisor and mentor, Dr. Harel Shouval for the continuous support and encouragement since I joined graduate school in fall 2009. In addition, I sincerely acknowledge Dr. Naveed Aslam for extending his MATLAB codes on which some of my initial work was based. I also with to thank Dr. Gastone Castellani and Dr. Yin Liu for painstakingly going over the manuscript and providing me with useful suggestions.

I wish to thank my fellow graduate students at University of Texas and my friends at Rice University and Baylor College of Medicine, who made my stay in Houston a pleasurable experience. Lastly, I would like to extend sincere gratitude to my friends and family without whose support this work would not have been possible.

I would to like to extend unrestricted access to all my models and its implementation to Dr. Shouval and his research group for future research.

Stochastic modeling and simulation of biochemical reaction kinetics

Animesh Agarwal, MSE
The University of Texas at Austin, 2011

Supervisor: Harel Shouval

Abstract: Biochemical reactions make up most of the activity in a cell. There is inherent stochasticity in the kinetic behavior of biochemical reactions which in turn governs the fate of various cellular processes. In this work, the precision of a method for dimensionality reduction for stochastic modeling of biochemical reactions is evaluated. Further, a method of stochastic simulation of reaction kinetics is implemented in case of a specific biochemical network involved in maintenance of long-term potentiation (LTP), the basic substrate for learning and memory formation. The dimensionality reduction method diverges significantly from a full stochastic model in prediction the variance of the fluctuations. The application of the stochastic simulation method to LTP modeling was used to find qualitative dependence of stochastic fluctuations on reaction volume and model parameters.

Table of Contents

List of Figures

Chapter 1

Introduction

Various internal functions of a cell occur in concert with each other. Successful functioning of the cell requires precise control of the cellular process. This control is often achieved through interaction of protein networks. Progress in molecular biology has led to identification of complex biochemical reaction networks involved in normal functioning of a cell. Defects in the functioning of these networks are often associated with diseases. Better understanding of the reactions involved in these biochemical networks provide us with an opportunity to build mechanistic mathematical models to predict the over all behavior of the network [1]. Such predictions ultimately help provide possible scientific explanations for observations made experimentally.

Mechanistic modeling of even a moderately large biochemical network can be quite a complex high dimensional problem and simulating such networks can be numerically challenging and time consuming [2]. Thus it is worthwhile to explore methods for dimensionality reduction. Molecular interactions are inherently stochastic because they are determined by collisions and bimolecular interactions that depend on thermal energy and on the orientations of the interacting molecules.

In this thesis, I deal with two questions concerning the stochastic simulation of reaction kinetics of protein network. In Chapter 2, I test the accuracy of one of the methods for dimensionality reduction, quasi-steady state approximation (QSSA) under a stochastic simulation framework. I test the QSSA based reduction method first on a canonical single enzyme reaction scheme which is used to obtain the Michaelis-Menton kinetic scheme. Then I extend the test of QSSA based dimensionality reduction to a basic post-transational bistable loop [3] that may be responsible for a basic substrate of learning and memory in the brain. In Chapter 3 of this thesis, I develop a reaction scheme that serves as a model for a translation based bistable loop that controls the long term maintenance of synaptic plasticity that may be implicated in long term-memory and learning. I then evaluate the sensitivity of the stochastic fluctuations on the parameters of the model.

One of the approaches to mechanistic modeling, mass action approach, assumes that molecular reactions can be characterized by the concentrations of the different molecular species. This approach describes the dynamics of a molecular network via a set of coupled ordinary differential equations. Such an approach is usually a good approximation when the molecular system is well mixed, and when the number of molecules is large, so that relative fluctuations from the mean are typically small.

Many biological systems have a small number of molecules that determine cell function, among them are gene expression processes that govern cell fate [4]. Processes such as transcription are outside the scope of this thesis

where we are primarily considering translational and post-translational processes. One of the biological systems in which stochastic fluctuations play a crucial role is synaptic plasticity - the molecular substrate of of learning and memory. Synaptic strengths between neurons is modulated by enzymatic reactions taking place in the small volume of a neuronal spine, a volume much smaller than that of a typical cell. At such a small volume, a small number of molecules are involved in the reaction such that the stochastic fluctuations may be responsible for determining the synaptic state. Various studies have examined the impact of such stochastic fluctuations during the induction [5, 6, 7, 8, 9, 10] and maintenance phases of synaptic plasticity [11, 12]. Stochastic fluctuations might be of importance not only when the actual sizes of compartments are small, but also when systems are not well mixed and as a result, have effective micro-domains in which relative fluctuations are significant.

There is a long history of methods for reducing the dimensionality of biochemical reaction networks. The set of approximations called the QSSA is based on the assumption that some of the molecular reactions are fast, and can therefore be assumed to reach quickly to steady state, and therefore some of the dynamical variables can be eliminated by replacing them with their steady state values [2]. Different variants of this approach have been proposed over the years [13].

For stochastic systems not only the mean of the molecular concentrations must be identified, but also the fluctuations around the mean, or more

generally, the distributions of the molecule numbers of the different molecular species. The dynamics of these distributions are characterized by the so called chemical master equation [14]. Usually the only practical way to estimate the joint distributions of the different chemical species is by using numerical Monte-Carlo methods. The most commonly used method is the exact stochastic simulation method proposed by Gillespie [15]. However, in complex networks this method is likely to be very slow. This is especially the case when molecular networks include a mix of slow and fast reactions. While the fast reaction consume most of the computational power, the slow reactions often determine the primary shape of the dynamics, and the magnitude of the stochastic fluctuations. For example, in case of the simple enzyme catalyzed reaction, the exact stochastic simulation algorithm can take a lot of time calculating the binding and unbinding of the substrate to the enzyme while we may really be interested in seeing the product formation dynamics [16]. Therefore, it is appealing to try and reduce dimensionality, using the QSSA in order to simplify the stochastic molecualr networks, and speed up the calculations of the distributions. Such an approach was proposed by Rao and Arkin (2003), who showed that such an approximation can produce a reasonable approximation of the dynamics of the mean concentration . Also, dimensionality reduction can lead to more intuitive expressions and possibly analytical solutions. So the relevance of this work goes farther than just improvement in speed of calculations.

The focus of Chapter 2 is whether the QSSA approach can approxi-

mate well not only the means of the distributions of the dynamical variables, but also their fluctuations, or more generally the complete distributions. We have tested this is two different molecular networks over a wide range of parameters. We have examined two examples; a simple enzyme reaction with a basal rate of reversal of product back to substrate and a kinase phosphatase bistable switch similar to the one suggested by Lisman in 1985 [3]. In our analysis, we have found several parameter sets for which a QSSA based reduction scheme fails to capture the distribution of substrate concentration around the mean as obtained from the full model. Also, we found no significant relationship between accuracy of a QSSA based reduction method in predicting the deterministic transients and the accuracy of such a reduction method in predicting the stochastic fluctuation at steady state. For the bistable reaction system, we found that a QSSA based reduction method made significant errors in prediction of residence times in each stable state by about two orders of magnitude.

An application of the stochastic simulation approach developed in this work is in the biochemical reactions controlling synaptic strength between neurons. Modeling and simulation of this biochemical reaction network has implications on understanding the mechanism of learning and memory. Learning and memory generally are persistent for a long time, even as long as the human lifetime. The basic substrate of learning and memory in the brain is synaptic plasticity and in particular long-term potentiation (LTP) of the synaptic efficacies. LTP has a late protein synthesis dependent phase (L-LTP)

that can last for many hours in slices, or even days in vivo [18, 19]. L-LTP is modulated by activity of certain proteins in the post-synaptic spine. It is interesting, that the proteins that are responsible for such long term maintenance of potentiation themselves have a turnover time ranging from only few minutes to few hours. Thus the protein concentration needed for L-LTP must be maintained by replenishment of proteins. Such a replenishment can occur either through diffusion from outside the spine or through translation of new protein molecules inside the spine. Here, we concentrate primarily on replenishment through translation of new protein molecules inside the spine.

The synaptic strength, and the equilibrium concentration of the protein molecules in turn, is specific to a synapse. Thus transcriptional control in the nucleus of the cell, that modulates the amount of mRNA at all synapses non-specifically, is insufficient to achieve synapse specific modulation of the protein concentrations.

Previous studies have shown that maintenance of L-LTP and memory can be accounted by persistent regulation of on-site synthesis of plasticity-related proteins by a self-sustaining regulation of translation. Such a system can exists as a bistable switch which is modulated by a second messenger based signal. An example is the αCaMKII -CPEB1 molecular pair that can act as a bistable switch with different total amounts of αCaMKII in potentiated and non-potentiated synapses [20].

The previous models for bistable synaptic efficacies, based on a translation level loop, were deterministic [20]. However, at the small volume of the

spine, the stochastic nature of chemical reactions might become very important and may cause a reversal of LTP. To test this, we look at the stochastic behavior of this system at very low poly-ribosome and mRNA concentration levels using the stochastic simulation algorithm (SSA) suggested by Gillespie [15]. To implement the stochastic simulation algorithm, we suggest a model of translation with explicit implementation of mRNA and poly-ribosome concentration in the post-synaptic spine.

In Chapter 3 of this thesis, we have examined the impact of stochastic fluctuations on length of maintenance of L-LTP. Specifically, we have tested the effect of simulation volume on the nature of stochastic fluctuations and found that L-LTP was stable in the upstate for greater than 30 hours for a reaction volume close to the volume of the spine. We have also tested the sensitivity of stochastic fluctuations to reaction rate parameters. In order to carry out these simulations we had to make simplifying assumptions about the translational mechanisms, and model them as an elementary kinetic step. In Chapter 3 and the appendix we examine the validity of this approximation as well.

Chapter 2

Precision of dimensionality reduction methods

Many biochemical networks have complex multidimensional dynamics. There is a long history of methods that have been used for dimensionality reduction for multidimensional reaction networks. In this chapter, we evaluate the applicability of one such method, the QSSA method, for dimensionality reduction in case of stochastic simulations. In the first section, the applicability of QSSA approach is evaluated for a canonical enzyme reaction kinetics. In the second section, we have extended the analysis to a well-known biochemical network model for the basic substrate of learning and memory as proposed by Lisman in 1985 [3]. Lisman's model is an example of bistable system. Our analysis has shown that a QSSA based dimensionality reduction method fails to match prediction of frequency of state transitions as predicted by a full-dimensional analysis.

2.1 Enzyme reactions

The quasi steady state approximation has been applied to a standard two-stage enzyme mediated catalysis reaction to obtain the Michaelis-Menton kinetics. Consider the following set of reactions.

$$
\begin{aligned}
E + S \quad &\underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \quad C \overset{k_2}{\longrightarrow} E + P \\
P \quad &\overset{k_3}{\longrightarrow} \quad S
\end{aligned}
$$

The first reaction indicates the reversible binding of the enzyme (E) and substrate (S) to form the enzyme-substrate complex (C). The enzyme-substrate complex either dissociates back to free enzyme and free substrate or a product molecule (P) is formed and released from the enzyme. The second reaction represents an irreversible conversion of products back to substrate via a separate independent process. Such an additional process is of interest as in several large networks of enzymatic reactions, individual dynamics of a particular substrate may strongly be affected by a background rate of conversion of products back to substrate. The addition of the second reaction also ensures that, at equilibrium, the substrate concentration is not equal to zero. This is especially important because we are interested in the distribution around steady state for stochastic simulations and such a distribution exists only for a non-zero steady state.

Using a mass action approach, a 3-dimensional system of differential equations describes the kinetics of the full model

$$
\begin{aligned}
\frac{\mathrm{d}S}{\mathrm{d}t} &= -k_1 E \cdot S + k_{-1} C + k_3 P \qquad (2.1) \\
\frac{\mathrm{d}C}{\mathrm{d}t} &= k_1 E \cdot S - (k_{-1} + k_2) C \qquad (2.2)
\end{aligned}
$$

$$\frac{\mathrm{d}P}{\mathrm{d}t} = k_2 C - k_3 P \tag{2.3}$$

$$E_T = E + C \tag{2.4}$$

The initial conditions can be chosen such that the mass conservation is not violated and all the concentrations are positive. One such set of initial conditions could be $S(0) = S_T, C(0) = 0, P(0) = 0, E(0) = E_T$ where $S_T = S + C + P$. These initial conditions are changed to match the initial conditions as required by the mass balance in the reduced model discussed below.

The QSSA assumes that at the time scale at which the substrate is being consumed (or product is being formed) by the first reaction, the concentration of enzyme-substrate complex is essentially not changing. Hence, the dimensionality of the set of differential equations, that govern the kinetics of the first reaction, can be reduced, by setting the time derivative of the enzyme-substrate complex to zero. The QSSA results in a 1-dimensional ODE that describes the kinetics of the reduced model [2]

$$C^*(t) = \frac{E_T \cdot S(t)}{k_m + S(t)} \tag{2.5}$$

$$P(t) = S_T - S(t) - C^*(t) \tag{2.6}$$

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -k_1 \left(E_T - C^*(t)\right) S(t) + k_{-1} C^*(t) + k_3 P(t) \tag{2.7}$$

where $k_m = (k_{-1} + k_2)/k_1$

Since $C^*(t)$ is now a function of $S(t)$, the initial conditions of Equations 2.1 - 2.4 that are consistent with QSSA are now subject to a more severe

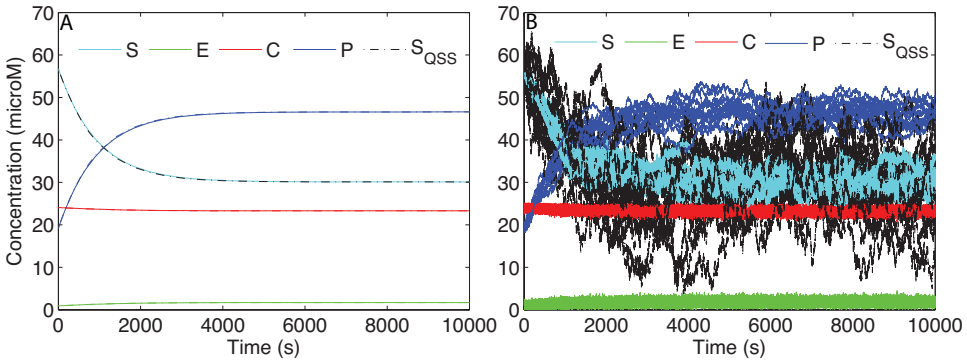

Figure 2.1: **A**. Quasi-steady state approximation. Comparison of deterministic solution of the complete enzymatic system and the reduced system with the quasi-steady state assumption. The dashed line is the substrate trace for the 1-d equation with quasi-steady state assumption. For the chosen parameters that satisfy the QSSA validity criterion, the dashed line is reasonably close to the substrate trace for the full model. The substrate trace for both the full and the reduced model settle down to the same steady state.(Parameters: $k_1 = 0.01, k_{-1} = 0.02, k_2 = 0.002, k_3 = 0.001, E_T = 25\mu M, S_T = 100\mu M$) **B**. Stochastic simulation of the chemical master equation for the complete set enzymatic reactions using the Gillespie algorithm. Overlaid on top, is the reduced stochastic simulation by implementing a Gillespie-like algorithm for the reduced chemical master equation, using the quasi-steady state assumption (Eqn 2.17). The mean of the stochastic simulation is well approximated by the deterministic simulation. The distribution around mean for substrate concentration for the reduced stochastic simulation is different than that for the full stochastic distribution. The volume of the reaction mixture is 10^{-17} L

restriction. The maximum allowed value of S is set at the initial condition by solving the mass balance equation: $S_T = S(0) + C^*(0) + P(0)$. The initial value of P can be chosen such that the mass conservation holds. For sake of convenience, we have chosen $P(0) = 0$. The initial conditions of the full model is also changed accordingly to match the reduced model.

Figure 2.1A (solid lines) shows the numerical solution of the system

ODEs for the full model. Overlaid on the same plot is also the numerical simulation for the 1-d ODE for time evolution of substrate in the reduced model (dashed line). In the example shown, the parameters are chosen such that the following condition for validity of QSSA is satisfied [13].

$$k_m + S_T \gg E_T \tag{2.8}$$

Hence the substrate trace for the reduced model is reasonably close to the substrate trace for the full model. Note that both the full model and the reduced model settle down to the same steady state.

The main interest in this work is not the deterministic solution to the ODEs governing the dynamics of the system. Instead, our focus here is the stochastic simulation. Formally, the joint distribution of the number of free substrate molecules (n_S) and number of complex molecules (n_C) in a stochastic system is given by the following chemical master equation:

$$\begin{aligned}
\frac{\mathrm{d}\rho(n_S, n_C)}{\mathrm{d}t} &= -\left((n_{E_T} - n_C)n_S k_1 + n_C(k_{-1} + k_2)\right)\rho(n_S, n_C) \\
&\quad -\left((n_{S_T} - n_C - n_S)k_3\right)\rho(n_S, n_C) \\
&\quad +(n_{E_T} - n_C + 1)(n_S + 1)k_1\rho(n_S + 1, n_C - 1) \\
&\quad +(n_C + 1)k_{-1}\rho(n_S - 1, n_C + 1) \\
&\quad +(n_C + 1)k_2\rho(n_S, n_C + 1) \\
&\quad +(n_{S_T} - n_S - n_C + 1)k_3\rho(n_S - 1, n_C)
\end{aligned} \tag{2.9}$$

Where n_{S_T} is the number of free product molecules plus the total number of S

molecules in bound and free form and n_{E_T} is the total number of E molecules in bound and free form.

It is usually not possible to obtain analytical solutions of the master equation. Instead, a numerical solution to Equation 2.9 can be obtained by implementing the stochastic simulation algorithm suggested by Gillespie [15]. Figure 2.1B shows the result of fifteen runs of Gillespies algorithm for the same set of parameters as that in Figure 2.1A. Note that the Gillespie algorithm simultaneously solves for the number of molecules of all the species in the reaction mixture. Also note that in Figure 2.1B, for comparison with Figure 2.1A, the output of the Gillespies algorithm has been suitably scaled with reaction volume, V and Avagadro's number, N_a to represent concentrations instead of numbers of molecules.

For the stochastic system the aim of QSSA is to reduce the dimensionality of the two dimensional master equation as described in Equation 2.9 to a one-dimensional master equation of the form:

$$\frac{\mathrm{d}\rho(n)}{\mathrm{d}t} = [r(n+1).\rho(n+1) - r(n).\rho(n)] + [g(n-1).\rho(n-1) - g(n).\rho(n)] \quad (2.10)$$

where $r(n)$ and $g(n)$ are the sink and source functions respectively. In such an equation, the sink function and the probability function must be subject to the following *boundary conditions*:

$$r(0) = 0 \quad (2.11)$$

$$\rho(i) = 0 \qquad \forall \quad i < 0 \quad (2.12)$$

In fact, for this kind of chemical master equation, there is a general recursive analytical solution of the form [14] (Also see Appendix A for proof):

$$\rho(n) = \frac{g(n-1)}{r(n)}\rho(n-1) \qquad \forall \quad n = 1, 2, 3 \ldots \infty \tag{2.13}$$

A method for reducing dimensionality of the system of chemical master equations for stochastic simulation (Eqn 2.9) was proposed by Rao et al (2003) , where it was suggested that the equation derived from law of mass action in a deterministic system (Eqn 2.7) can be used to directly write a reduced chemical master equation for stochastic evolution of the substrate molecule. Accordingly, the reduced chemical master equation is obtained as follows:

We first rewrite Eqn 2.7 in terms of number of molecules instead of concentrations by replacing concentration values (c) with molecule numbers (n) using the relationship $c = n/(N_a \cdot V)$ where , N_a is Avogadros number and V is the volume of the reaction mixture. The following equation is obtained:

$$n_{C^*} = \frac{n_{E_T} \cdot n_S}{k_m \cdot N_a \cdot V + n_S} \tag{2.14}$$

$$n_P = n_{S_T} - n_S - n_{C^*} \tag{2.15}$$

$$\frac{\mathrm{d}n_S}{\mathrm{d}t} = -\frac{k_1}{N_a \cdot V}\left(n_{E_T} - n_{C^*}\right) n_S + k_{-1} n_{C^*} + k_3 n_P \tag{2.16}$$

Subsequently the origin of each term in equation 2.16 is identified as a source or sink term, in order to write the following chemical master equation:

$$\begin{aligned}\frac{\mathrm{d}\rho(n_S)}{\mathrm{d}t} = & \quad [r(n_S+1).\rho(n_S+1) - r(n_S).\rho(n_S)] \\ & + [g(n_S-1).\rho(n_S-1) - g(n_S).\rho(n_S)]\end{aligned} \tag{2.17}$$

where sink and source functions are identified as follows:

$$r(n) = \frac{k_1}{N_a \cdot V}\left(n_{E_T} - n_{C^*}(n)\right) n \tag{2.18}$$

$$g(n) = k_{-1} n_{C^*}(n) + k_3 n_P(n) \tag{2.19}$$

where,

$$n_{C^*}(n) = \frac{n_{E_T} \cdot n}{k_m \cdot N_a \cdot V + n} \tag{2.20}$$

$$n_P(n) = n_{S_T} - n - n_{C^*}(n) \tag{2.21}$$

Algebraic simplification of Eq 2.16 can give rise to another possibility for the choice of $r(n)$ and $g(n)$ as follows:

$$r1(n) = \frac{V'_{max} n}{k'_m + n} \tag{2.22}$$

$$g1(n) = k_3(n_{S_T} - n) \tag{2.23}$$

where $V'_{max} = E_T(k_2 + k_3) \cdot N_a \cdot V$ and $k'_m = k_m \cdot N_a \cdot V$. There is no clear reason to prefer one choice of $r(n)$ and $g(n)$ over another, as long as the boundary conditions of the system are satisfied, any choice of $r(n)$ and $g(n)$ should suffice. It is interesting to note though that any reasonable choice of $r(n)$ and $g(n)$ preserves the mean of the stationary stochastic distribution, the specific choice however, significantly affects the variance around this mean.

Equation 2.17 can be solved either analytically or by using the numerical Gillespie algorithm [17]. In figure 2.1B the results of fifteen runs of the reduced stochastic simulation (black) are overlaid on the traces the full stochastic simulations (cyan). The reduced stochastic simulation in this figure is for the choice of r_n and g_n as in equation 2.18 - 2.19. It appears that the

mean of the stochastic simulations (Fig 2.1B) is quite close to the deterministic simulation (Fig 2.1A) for both the reduced and full stochastic simulation.

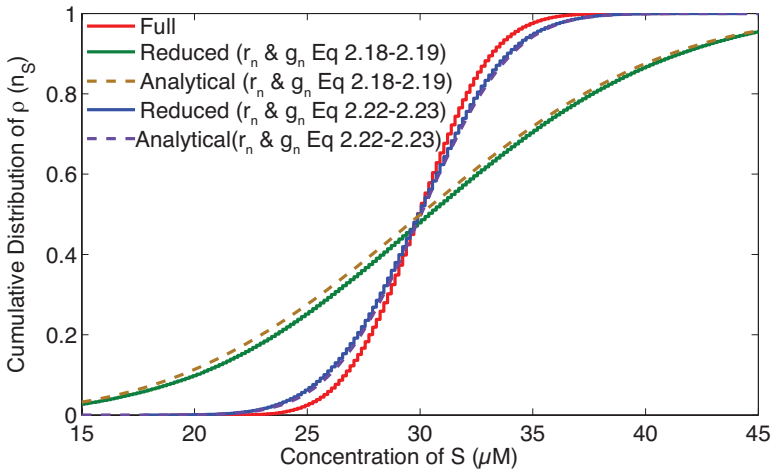

Figure 2.2: C-storeomparison of the cumulative distribution of probability of a certain concentration of the substrate in the reaction volume after the corresponding deterministic simulation has reached steady state. The cumulative distribution for the numerical solution of the chemical master equation for the complete reaction system has same mean (full: 29.92, reduced (Eq 2.18 and 2.19) : 30.66, reduced (Eq 2.22 and 2.23) : 29.88, reduced analytical (Eq 2.24): 30.12) as that of the reduced system but a different standard deviation (full: 2.60, reduced (Eq 2.18 and 2.19) : 8.36, reduced (Eq 2.22 and 2.23) : 3.21, reduced analytical (Eq 2.24): 3.16). The Kolmogorov Smirnov test shows that the three distributions are significantly different (p-value $= 1e-3$). The analytical and the numerical solution of the reduced chemical master equation (Eqn 2.17) yield identical results.

However, here we concentrate on the fluctuations of the distributions around the mean. To simplify this task we concentrate on the distributions at steady state. This is done by running the full and reduced model stochastic simulation for a long time and comparing the distribution of the full and reduced models at steady state. We use the part of the substrate trace where

the deterministic counterpart has already reached steady state ($t \geq 6000$ seconds in the example) and calculate the probability of a certain concentration of substrate when the system is in overall equilibrium. The cumulative distribution of this probability is plotted in Figure 2.2. The Kolmogorov Smirnov test shows that the cumulative distribution for the complete reaction system is significantly different from that for the reduced system.

Equation 2.13 can be used to write the analytical solution for the 1-D master equation (Eqn 2.17). For the choice of r_n and g_n in Eqn 2.22 and 2.23:

$$p_n = p_{n-1}\left[\frac{k_3(n_{S_T}-n+1)(k'_m+n)}{V'_{max}\cdot n}\right] \tag{2.24}$$

$$\Rightarrow p_n = \prod_{i=1}^{n}\frac{k_3\cdot(n_{S_T}-i+1)}{V'_{max}i}(k'_m+i)p_0 \tag{2.25}$$

$$= \left(\frac{k_3}{V'_{max}}\right)^n\binom{n_{S_T}}{n}\prod_{i=1}^{n}(k'_m+i)p_0 \tag{2.26}$$

$$= \left(\frac{k_3}{V'_{max}}\right)^n\binom{n_{S_T}}{n}k'^{\,(n)}_m p_0 \tag{2.27}$$

where $x^{(n)} = x(x+1)(x+2)\cdots(x+n-1) = \frac{(x+n-1)!}{(x-1)!}$ is the rising factorial symbols (or Pochhammer symbol) and p_0 is set by normalization. This analytical solution exactly matches the numerical estimates of the distribution of the reduced model (Fig 2.2). The ability to extract an analytical solution for the reduced model is another motivation for using the QSSA approach, because it sometimes allows us to obtain a closed form solution of the complete distribution of the system, and an understanding of how the different properties of the distribution depend on the system parameters.

The significant difference obtained in the distribution around mean

between the full stochastic model and the reduced model suggests that constructing a reduced chemical master equation from QSSA reduced deterministic equation may not always yield desired accuracy in stochastic simulations, even though the traditional condition for validity of quasi-steady state approximation (Eqn 2.8) is satisfied.

One could hypothesize that a good approximation of the deterministic transient would produce a good approximation of the distribution at steady state in stochastic simulation. To test this, we examined if the accuracy of the reduced deterministic model is related to the ability of the reduced stochastic model to estimate the distribution of the full stochastic model. In order to do this we measured the normalized mean square error between the deterministic transient from the full model and the reduced model and compared it to the Kullback-Leibler divergence between the steady state distributions of the full and reduced models (See Appendix). These results, which are displayed in figure 2.3, show no apparent relationship between accuracy of the QSSA deterministic solution and the closeness of stochastic distributions. Of particular interest is that even when the validity condition for the QSSA is observed (black symbols), the divergence between the reduced and full model distributions can be significant.

2.2 A bistable kinase-phosphatase molecular switch

Stochastic fluctuations have a significant impact in bistable systems, where these fluctuations determine the transitions between the two stable

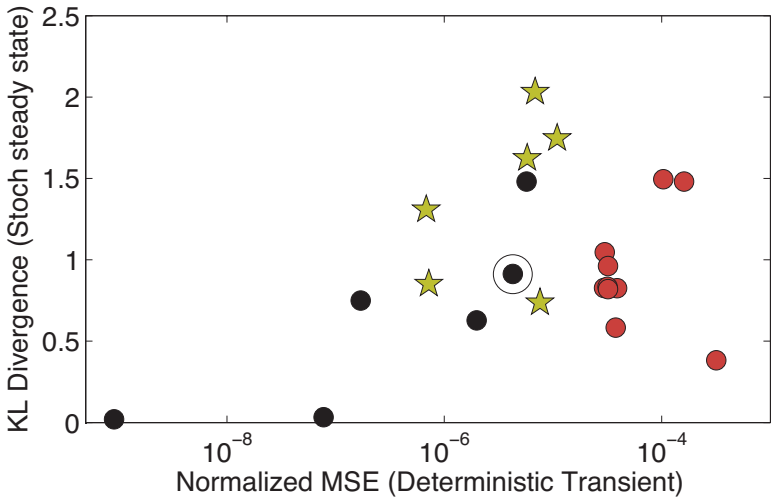

Figure 2.3: Relationship between the accuracy of the quasi-steady state approximation in predicting the deterministic transients and the accuracy of the reduced stochastic model to predict steady state distributions. The x axis represents the accuracy of the deterministic approximation using a normalized mean square error, and the y axis measures the Kullback-Liebler divergence between the distributions of the full and reduced models at steady state; a measure of the distance between these distributions. Each point represents a different set of parameters. The black points are the ones where the condition of validity of quasi-steady state approximation (Equation 2.8) is satisfied. Star-shaped points are for parameters consistent with Barik et. al. (2008) with various levels of k_3. The point with a circle around it is the parameter set that is used for simulations in Fig 2.1 and 2.2

points. Strictly speaking, a bistable system exists only in a deterministic paradigm where a system stays in one of the stable states indefinitely. The state that the system finally settles down in, depends on the boundary conditions. In a system that has stochastic fluctuations, however, in general the system does not stay in one stable state indefinitely. Irrespective of the boundary conditions, once the system has settled in one of the stable states,

the stochastic fluctuations cause transitions from one state to another.

A simple example of such bistable system was proposed by Lisman (1985) . The original Lisman's model did not explicitly describe the reaction scheme. We have postulated a reaction scheme which is qualitatively consistent with the original Lisman's model.

The reaction scheme for kinase-phosphatase molecular switch is as follows:

$$\begin{aligned} K + K_P &\underset{r_{-1}}{\overset{r_1}{\rightleftarrows}} C_1 \overset{r_2}{\longrightarrow} 2K_P \\ P + K_P &\underset{r_{-3}}{\overset{r_3}{\rightleftarrows}} C_2 \overset{r_4}{\longrightarrow} K + P \\ K &\overset{r_5}{\longrightarrow} K_P \end{aligned}$$

The phosphorylated form of kinase molecule, K_P, is active while the dephosphorylated form of kinase. K is inactive. In the first reaction, an active kinase molecule phosphorylates an inactive kinase molecule. This process is called *auto-phosphorylation*. In the second reaction, an active kinase molecule is dephosphorylated by binding to a phosphatase molecule. The third reaction represents the basal level of phosphatase activity that converts an active kinase molecule to an inactive kinase molecule. This reaction is not in the original Lisman reaction scheme but we added it to avoid an absorbing state (i.e. the system gets caught in the lower state because all the concentrations are zero and not because of absence of stochasticity). Note that here the enzyme and the substrate are different states of the same molecular species; making it difficult to obtain a validity condition for the QSSA.

The 'full' deterministic model for the above set of reactions is obtained by applying the of law of mass action.

$$\frac{\mathrm{d}K_P}{\mathrm{d}t} = -r_1 K \cdot K_P + (r_{-1} + 2r_2)C_1 - r_3 K_P P + r_{-3} C_2 + r_5 K \quad (2.28)$$

$$\frac{\mathrm{d}C_1}{\mathrm{d}t} = r_1 K \cdot K_P - (r_{-1} + r_2)C_1 \quad (2.29)$$

$$\frac{\mathrm{d}C_2}{\mathrm{d}t} = r_3 P \cdot K_P - (r_{-3} + r_4)C_2 \quad (2.30)$$

$$K_T = K + K_P + 2C_1 + C_2 \quad (2.31)$$

$$P_T = P + C_2 \quad (2.32)$$

where K is the concentration of the free dephosphorylated kinase, K_P is the concentration of the free phosphorylated kinase, P is the concentration of free phosphatase, C_1 is the concentration of the kinase auto-phosphorylation complex and C_2 is the concentration of the kinase dephosphorylation complex. K_T and P_T are total kinase and total phosphatase concentrations respectively. The initial conditions are $C_1(0) = C_2(0) = 0$, $P(0) = P_T$, $K(0) = K_T$ for low steady state and $K_P(0) = K_T$ for high steady state. This set of equations has two stable equilibrium solutions. This means that for the same set of parameters, starting from various initial conditions will lead to one of two different stable steady states. Such a system of equations is said to represent a *bi-stable system*. Figure 2.4A (solid lines) shows the result of numerical simulation of Equations 2.28 - 2.32 for two intial conditions. When $K_P(0) = 25\mu M$, the steady state concentration of K_P is $18.4\mu M$ and when $K_P(0) = 4\mu M$, the steady state concentration of K_P is 0.8 μM.

Applying the quasi-steady state assumption ($\frac{dC_1}{dt} = 0$ and $\frac{dC_2}{dt} = 0$) to

Equations 2.28 - 2.32, we obtain a 1-dimentional ODE for the concentration of phosphorylated kinase, that describes the kinetics of the reduced model (Equation 2.33).

$$\begin{aligned}\frac{\mathrm{d}K_P}{\mathrm{d}t} &= \frac{r_1 r_2 K_P + r_5(r_{-1}+r_2)}{r_{-1}+r_2+2r_1K_P}\left[K_T - K_P - \frac{r_3 P_T K_P}{r_{-3}+r_4+r_3K_P}\right] \\ &\quad - \frac{r_3 r_4 K_P P_T}{r_{-3}+r_4+r_3K_P} \end{aligned} \tag{2.33}$$

This QSSA equation is now analogous to the Lisman's model though it quantitatively differs from it. Numerical simulations of Equation 2.33 are shown as dashed lines in figure 2.4A. Fixed points of the system can be obtained analytically by setting the left hand side of Equation 2.33 to zero.. Note that the fixed points (steady states) of the reduced 1-d system are identical to the fixed points of the full model of Equations 2.28 - 2.32.

In order to construct a chemical master equation, the origin of each term in Equation 2.33 is taken into account to define the source and sink terms as follows:

$$\begin{aligned}\text{sink} &= \frac{r_1 r_2 K_P + r_5(r_{-1}+r_2)}{r_{-1}+r_2+2r_1K_P}\left[K_T - K_P - \frac{r_3 P_T K_P}{r_{-3}+r_4+r_3K_P}\right] \\ \text{source} &= \frac{r_3 r_4 K_P P_T}{r_{-3}+r_4+r_3K_P}\end{aligned}$$

Note that here again, different possible choices of source and sink terms can be formulated keeping the steady state the same. Figure 2.4B shows the

source and sink terms plotted separately as a function of K_P. The points of intersection of the solid and the dashed lines are the fixed points of Equation 2.33. Figure 2.4C shows the potential energy wells for the two steady states for the given set of parameters (See Appendix). Stochastic fluctuations between the two stable steady states in the 1-d reduced model are seen as transitions between the the two energy minima 2.4C.

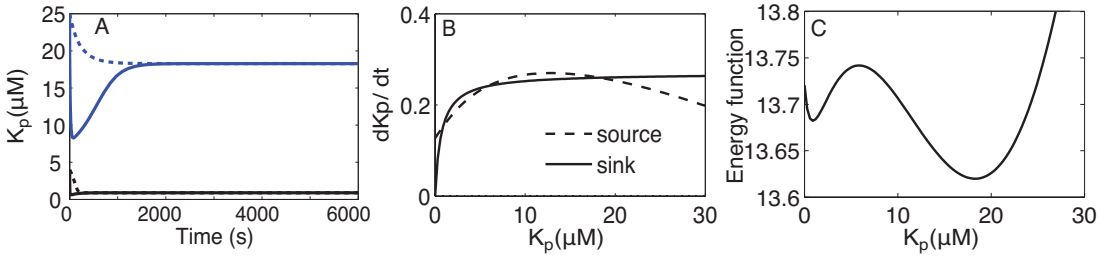

Figure 2.4: **A**. Deterministic simulation for the Kinase-Phosphatase switch showing bi-stability. The solid lines are obtained from the simulation of the full system. The dashed lines are from simulations with QSSA. Blue lines are for initial condition $K_P(0) = 25\mu M$ and black lines are for initial condition $K_P(0) = 4\mu M$
B. The source and sink terms for the quasi-steady state reduced model as a function of phosphorylated kinase concentration. The points of intersection of the solid and dashed lines are the fixed points of the system.
C. Potential energy wells for the fixed points in the reduced model. Note that the fixed points of the reduced model are identical to the fixed points of the full model. (Parameters: $r_1 = 0.001, r_{-1} = 0.002, r_2 = 0.02, r_3 = 0.08, r_{-3} = 0.001, r_4 = 0.0539, r_5 = 0.00212, K_T = 60\mu M, P_T = 5\mu M$)

Stochastic simulations for the kinase phosphatase switch are carried out by implementing the Gillespie's algorithm for the reaction scheme described above. Figure 2.5A shows a stochastic simulation of the complete reaction system. The solid lines represent the fixed point of the system. Note that

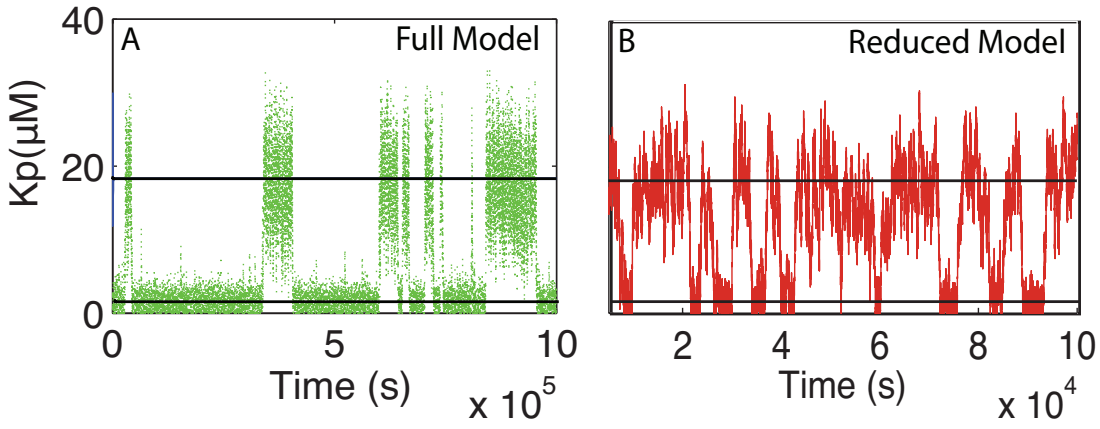

Figure 2.5: Stochastic switching between bistable states. **A**. Green line is the stochastic simulation of the complete reaction system using the Gillespie algorithm. The initial conditions are set to the upper steady state value. The black horizontal line indicates the two stable steady state concentration levels for each species. Note that the mean of the stochastic fluctuations within an equilibrium state is close to the steady state concentration. **B**. The red line is the stochastic simulation using the reduced master equation and a Gillespie-like algorithm. Note the difference in the scale of x-axis in the plots on the left and right.

the mean of the stochastic fluctuations in a given stochastic equilibrium well approximates the deterministic steady state (solid lines). Also note that a large fluctuation is enough to tip the system into the other stable state. This switching of states is very sensitive to the magnitude of stochastic fluctuations around the mean. To reduce the dimensionality of the stochastic simulation, a scheme similar to Equations 2.16 - 2.19 is applied. Equation 2.33 written in terms of number of molecules instead of concentration is:

$$\frac{\mathrm{d}n_{K_P}}{\mathrm{d}t} = \frac{V_{max_1} n_{K_P} + r_5 k_{max_1}}{k_{max_1} + n_{K_P}} \left[n_{K_T} - n_{K_P} - \frac{V_{max_2} n_{K_P}}{k_{max_2} + n_{K_P}} \right]$$

$$- \quad \frac{V_{max_3} n_{K_P}}{k_{max_2} + n_{K_P}} \tag{2.34}$$

where,

$$V_{max_1} = \frac{1}{2} r_2, \qquad V_{max_2} = n_{P_T}, \qquad V_{max_3} = r_4 n_{P_T}$$
$$k_{max_1} = \frac{1}{2} \frac{r_{-1} + r_2}{r_1 / N_a V}, \qquad k_{max_2} = \frac{r_{-3} + r_4}{r_3 / N_a V}$$

Similar to Equation 2.17 - 2.19, the following one dimensional chemical master equation is obtained:

$$\begin{aligned} \frac{\mathrm{d}\rho(n_{K_P})}{\mathrm{d}t} &= [r(n_{K_P}+1).\rho(n_{K_P}+1) - r(n_{K_P}).\rho(n_{K_P})] \\ &\quad + [g(n_{K_P}-1).\rho(n_{K_P}-1) - g(n_{K_P}).\rho(n_{K_P})] \end{aligned} \tag{2.35}$$

where

$$r(n) = \frac{V_{max_3} n}{k_{max_2} + n} \tag{2.36}$$

$$g(n) = \frac{V_{max_1} n + r_5 k_{max_1}}{k_{max_1} + n} \left[n_{K_T} - n - \frac{V_{max_2} n}{k_{max_2} + n} \right] \tag{2.37}$$

Note that here too, the master equation is not unique. Equation 2.35 is solved using a Gillespie like algorithm. Figure 2.5B shows the results of the numerical simulation of Equation 2.35. Notice that the switching is much faster (Note the different time scale). This is because the variance of K_P is larger in case of the reduced stochastic simulation. Figure 2.6 shows the stationary distribution of the probability density of a particular concentration of K_P for the full model and the reduced model. An analytical expression for such a stationary distribution is also obtained using Eqn 2.13. Notice that

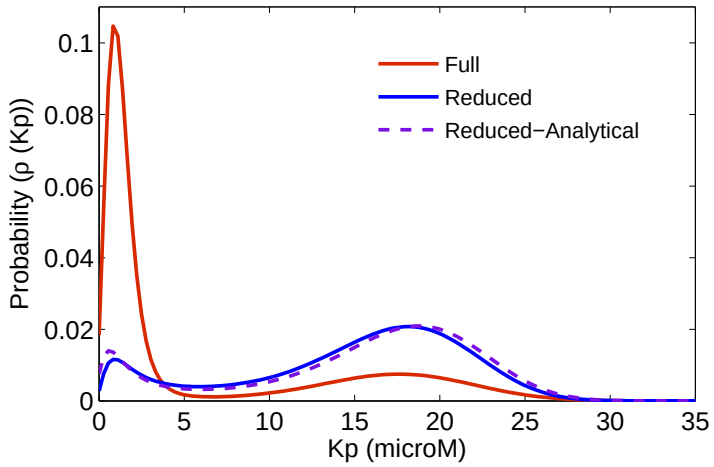

Figure 2.6: Histogram showing difference in upstate and downstate distribution for the full model and quasi steady state model for stochastic simulations. $Kp > 5.7\mu$ M is classified as upstate and $Kp < 5.7\mu$ M is classified as downstate.

the probability density functions for the reduced model is very different from that for the full model. The bimodal representation in the two cases is due to bistable nature of the system. Another statistic that characterizes this system is the characteristic dwell time at or near each of the stable states. To do this we set an upstate threshold and a downstate threshold. We count the time after the system goes below the downstate threshold until it hits the upstate threshold as 'downstate residence time' and the time after the system goes above the upstate threshold until it hits the downstate threshold as 'upstate residence time'. We set the same threshold for the full and the reduced model. For a simulation run over a long time (5×10^8 seconds), there are many transitions between the states, with a different residence time for each

transition; resulting in a distribution of the residence times. In our observation, we found that mean of both 'upstate residence time' and 'downstate residence time' were significantly different between the full and reduced model (Fig 2.7) (Kolmogorov-Smirnov test, p-value $= 1e - 5$). Specifically, the full model predicts much longer residence times in both the 'up' and 'down' states. This trend is consistent with the results in Fig 2.5 that show a smaller level of fluctuations in the full model, since these fluctuations drive the transitions between the states.

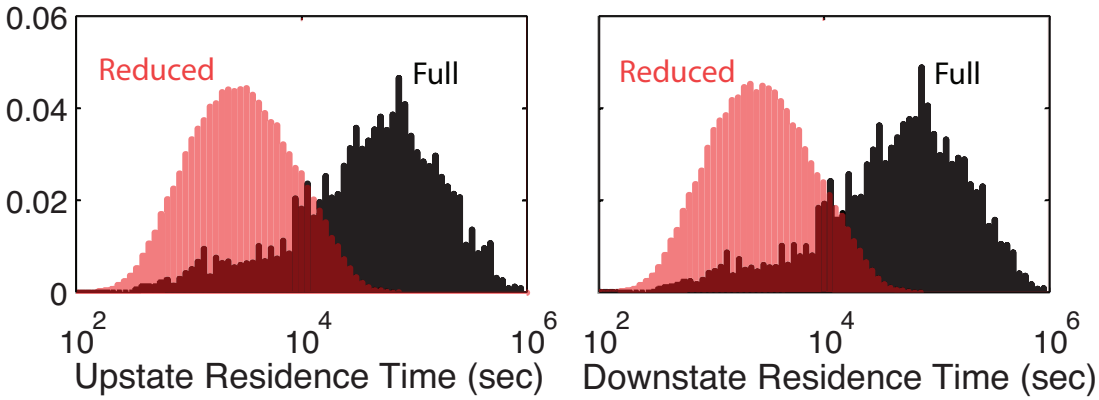

Figure 2.7: Distribution of residence time in the upstate and downstate for the full model and QSS model. The total simulation time over which this histogram is calculated is 5×10^8 seconds. The mean upstate residence time for full model is 8.34×10^4 seconds while that for reduced model is 4.92×10^3 seconds. The mean downstate residence time for full model is 9.30×10^4 seconds while that for reduced model is 4.56×10^3 seconds.

Chapter 3

Stochastic fluctuations in protein synthesis

Certain cellular processes require maintenance of stable protein concentration in a certain region. In presence of protein turnover, constant replenishment of protein molecules is required. This replenishment can either be provided by flow of molecules from outside the region or in-situ synthesis of proteins via translation. In the case of long term maintenance of synaptic plasticity, concentration of certain plasticity related proteins are locally modulated through a translation based positive feedback loop. The process of translation has inherent stochastic fluctuations that arise because of the small number of polyribosome complexes that act as sites for translation.

In this chapter, we have considered the stochastic fluctuations in the synthesis of a plasticity related protein, αCaMKII. The molecular control mechanism for the concentration of αCaMKII might consist of autophosphorylation loop linked with an upregulation of local synthesis that might exist as a bistable system. We have used stochastic simulation methods to evaluate the stability characteristics of a system that exhibits bistability under deterministic analysis. We have also evaluated the impact of choice of parameters of the reaction model on the stability characteristics.

3.1 Biochemical network model

The biochemical network responsible for maintaining the late phase of LTP might consist of a self sustaining autophosphorylation kinase loop interacting with a post-transcriptional up-regulation of protein synthesis [3]. An example of such an interaction network is the Calcium Calmoduling dependent αCam-KinaseII autophosphorylation coupled with the up-regulation of local αCam-KinaseII translation via CPEB1 based activation of its mRNA. In our previous study [20], the molecular interaction model comprised of an inactive, dephosphorylated, an inactive, phosphorylated and an active, phosphorylated form of αCaMKII and an inactive and active form of CPEB1. The bistability was due to interaction of Ca^{2+}-Calmodulin, auto-phosphorylation activation, spontaneous degradation and synthesis of αCaMKII. This model could successfully account for maintenance of L-LTP over a long period of time and also proposed an explanation for why application of protein synthesis and αCaMKII inhibitors at induction and maintenance phases of L-LTP result in very different outcomes [20, 21, 22]. The details of this protein interaction network can be found in [20]. In our current model, we have condensed the two step phosphorylation mechanism for αCaMKII to a single step. Also, we have assumed that activated CPEB1 activates mRNA which then binds preferentially to poly-ribosome, as compared to a non-active mRNA, for αCaMKII synthesis. We show that this system can also act as a bistable switch.

The reaction scheme shown in figure 3.1 can be modeled as a set of ODEs using the law of mass action approach. There are several free param-

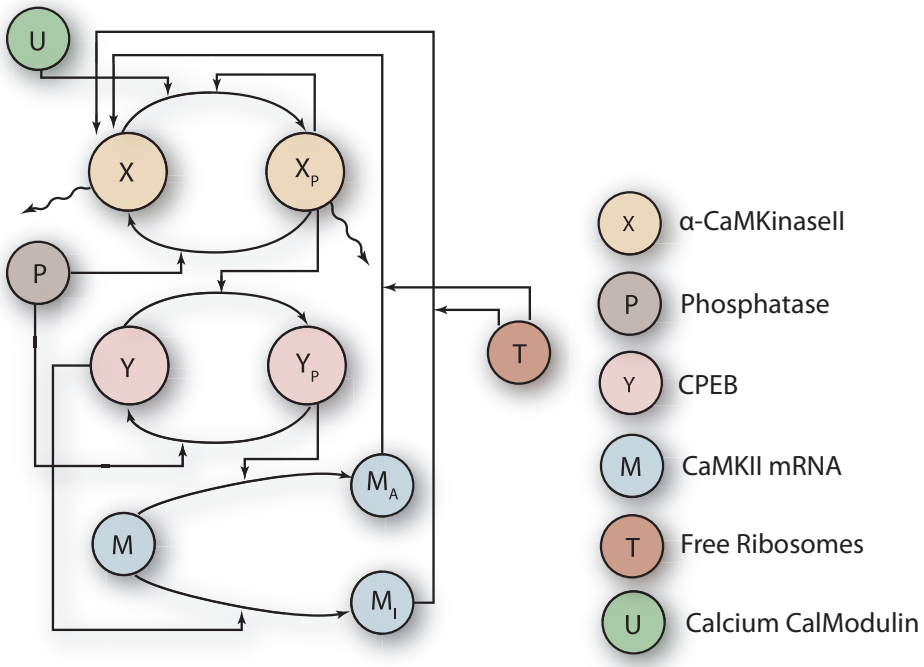

Figure 3.1: A schematic showing the reaction scheme in the model. Every circle is a component. An arrow starting from one circle and ending at another indicates a substrate product relationship. An arrow starting from a circle and ending at an arrow indicates that component is an enzyme for that reaction. The wavy arrows represent spontaneous turnover.

eters in the model that have been adjusted to extract bistable behavior from the system. Bistability in this case means that, starting from different initial conditions leads to one of two different equilibrium conditions. The two equilibrium conditions are differentiated from each other in the total concentration of CaMKII molecules (free dephosphorylated, free phosphorylated and that in complexes). Results of a simulation showing bistability are shown in figure 3.2. In this figure we see two dynamic deterministic simulations that differ

in initial conditions. Starting from the same total αCaMKII concentration, if all of CPEB1 is phosphorylated initially, the total αCaMKII concentration at equilibrium is 0.14 μM (downstate) and if all of CPEB1 is dephosphorylated initially, the total αCaMKII concentration at equilibrium is 5.4 μM (upstate). The inset shows the transients in the first 10 hours. A critical model parameter, the rate of protein production, has been adjusted to that observed experimentally [23].

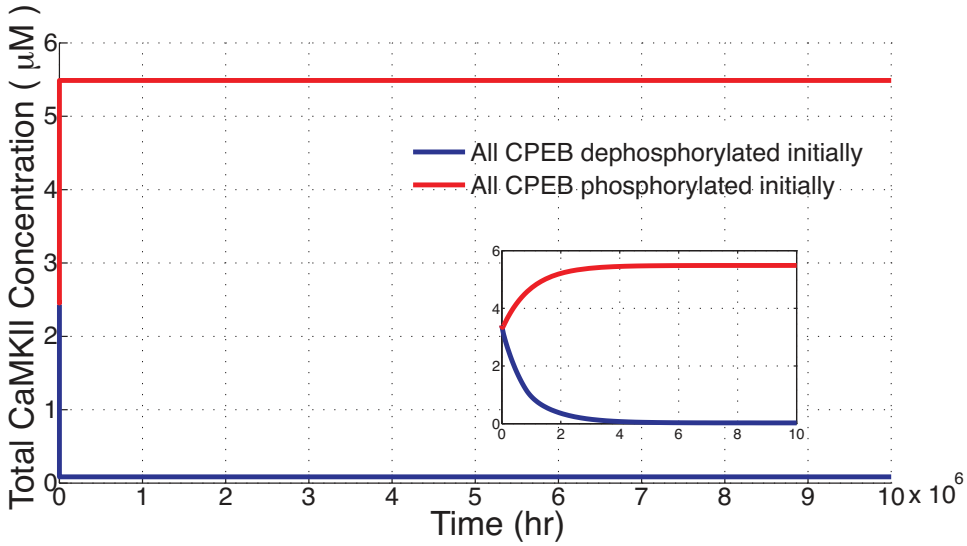

Figure 3.2: Deterministic simulation showing bistability. Starting with all the CPEB molecules in the dephosphorylated state leads to a 'downstate' and starting with all the CPEB molecules in the phosphorylated state leads to an 'upstate'. The inset shows the initial transient to upstate and downstate.

3.2 Stochastic simulations

A deterministic treatment of the reaction model provides an explanation for the experimentally observed behavior of the late phase of LTP. However, the chemical reactions involved in the model have an inherent stochasticity associated with them. There are two main reasons due to which stochastic fluctuations can play a crucial role in determining the system dynamics. First, these reactions take place in the small volume of the post-synaptic spine; at such low reaction volumes, the number of interacting molecules is very small (in the order of few tens to few hundreds). Second, there are a small number of polyribosome complexes in the spine, thus production of proteins is essentially stochastic in nature.

Large stochastic fluctuations can destabilize the the steady states and induce rapid transitions between these two states such that the total CaMKII concentration falls to the level of downstate during the maintenance phase of L-LTP thereby causing a reversal of LTP. These transitions might be approximately symmetric between the states. Thus, correctly characterizing the stochastic fluctuations is important for deriving predictions about the maintenance of L-LTP from a translation based model.

In order to test this we carried out stochastic simulations using Gillespie's SSA. This algorithm is valid when each reaction being modeled is essentially an elementary reaction step. An elementary reaction is a reaction for which no reaction intermediates have been detected or need to be postulated in order to describe the chemical reaction on a molecular scale. An elemen-

tary reaction is assumed to occur in a single step and to pass through a single transition state [24]. In case of a bimolecular elementary reaction, the product is formed only when the collision has the right orientation and enough energy to cross the reaction activation barrier. Such elementary reactions are characterized as Poisson processes with an exponential distribution of the interval between successive product formation. The translation process, as modeled here, is technically not an elementary reaction step as it is made up of several elementary steps. In Appendix B, we provide some motivation as to why the use of the SSA is justified in case of translation when several polyribosome complexes involved in protein formation are out of phase with each other.

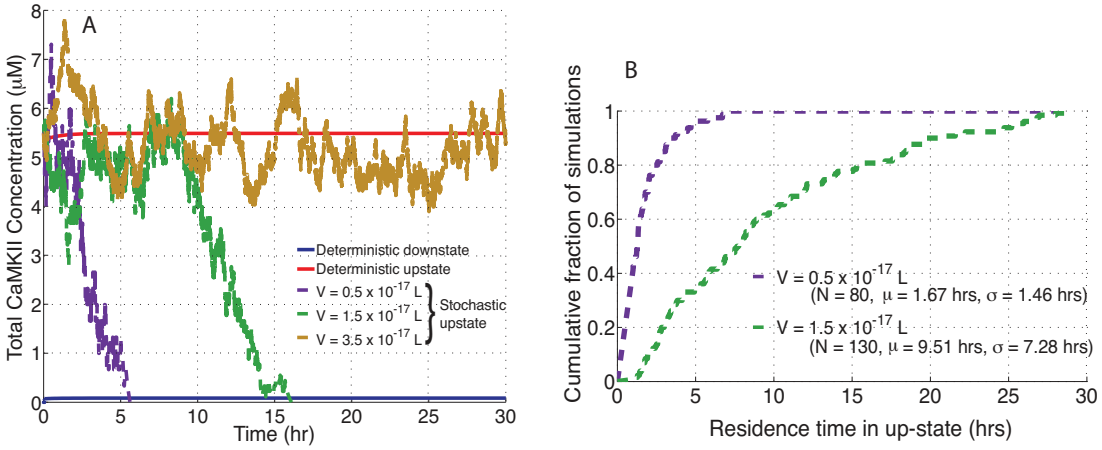

Figure 3.3: **A** Single stochastic run for three different reaction volumes. The overall fluctuation level is higher for smaller volume which destabilizes the upstate. **B** Cummulative distribution of residence time in upstate. For $V = 0.5 \times 10^{-17}$ L, the cumulative distribution if obtained from 80 independent runs, all starting from upstate. For $V = 1.5 \times 10^{-17}$ L, the cumulative distribution if obtained from 130 independent runs, all starting from upstate.

In this thesis, we have tested the effect of reaction volume and the choice

of free parameters on the level of stochasticity. To test the effect of volume, we carried out the simulations at three different reaction volumes. Figure 3.3A shows single runs at the three different reaction volumes. Figure 3.3B shows the cumulative distribution of residence time in the upstate for two different reaction volumes. All the simulations are started at the deterministic up-state for total concentration of CamKII. For a reaction volume of 0.5×10^{-17} L, which is 1/25th the volume of a typical spine (about 12.5×10^{-17} L), the total CamKII concentration reverts to down state in a mean time of 1.67 hours. For a reaction volume of 1.5×10^{-17} L, which is 1/8th the volume of a typical spine, the mean time of reversal of upstate is 9.51 hours. For a reaction volume of 3.5×10^{-17} L, which is 1/3rd the volume of a typical spine, the upstate was stable for longer than 30 hours in several runs. In fact we were never able to observe a transition from upstate to downstate for the this reaction volume. Using the Komogorov-Smirnov test, we find that the cumulative distribution of residence time in upstate is significantly different for the two reaction volumes where we observed transition up to almost 100% confidence.

Many of the parameters in this model are not based on experimental observations. This is because the rate constants for elementary reaction steps used in this model are very difficult to be measured in experiment. To test the effect of choice of free parameters on the level of stochasticity, we multiplied both the forward and backward rate constants for binding of mRNA to active CPEB molecule by alpha. This ensured that the deterministic fixed points remained the same. For an alpha less than one, the binding reaction is slower

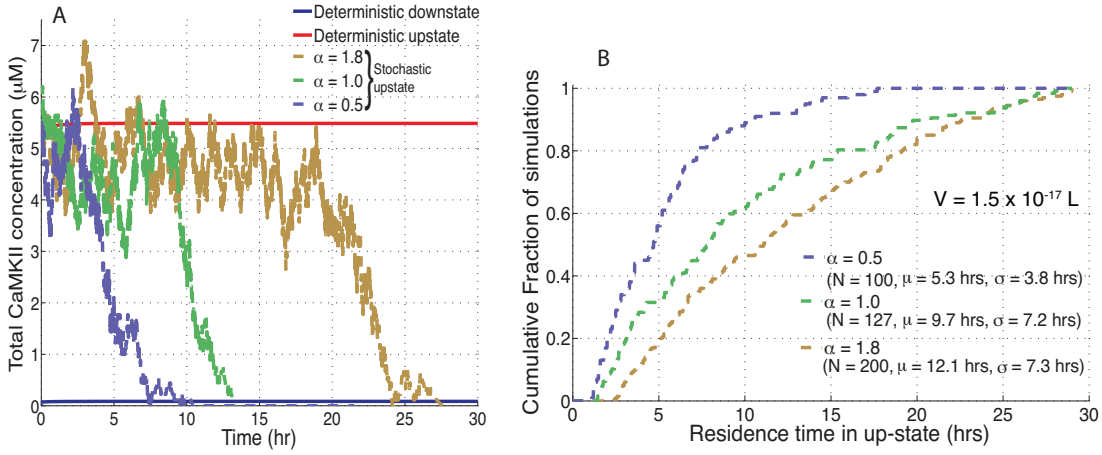

Figure 3.4: **A** Single stochastic run for three different choices of free parameters while keeping the deterministic fixed points the same. The overall fluctuation level changes with the choice of free parameters. The reaction volume in these simulations is 1.5×10^{-17} L. **B** Cummulative distribution of residence time in the upstate.

than when alpha is one. Slower reaction means greater stochastic fluctuations. Similarly, for alpha greater than one, the reaction is faster and there is smaller stochastic fluctuations. The results for single runs with three different values of alpha are shown in Figure 3.4A. As the reaction is made faster by changing the choice of free parameters while keeping the deterministic fixed points unaltered, the level of stochasticity decreases and the upstate is stable for a longer period of time. Figure 3.4B shows the cumulative distribution of residence times in upstate for $\alpha = 0.5$, 1.0 and 1.8 for 100, 127 and 200 independent simulations respectively. We see that the mean residence time in the upstate increases as the binding of CPEB1 to mRNA is made faster by increasing α. Using the Komogorov-Smirnov test, we find that the difference between all

three distributions is statistically significant (p-value $= 1e-2$). Note that the forward and backward rate constants for this reaction are only chosen as an example. A similar test with any other parameters of the model shall yield similar results.

Chapter 4

Discussion

The cellular processes that characterize the life of a cell are underlined by biochemical reactions. The understanding of dynamic and equilibrium behavior of these biochemical reactions are crucial to understanding the processes in a cell. A mis-function in the biochemical reactions often leads to diseases. So an understanding of these reactions are also often essential for designing drugs to treat diseases [1].

The kinetics of biochemical reactions have been studied for a long time. Recent advances in understanding of the elementary nature of these reactions have made possible detailed models for the kinetics of such reactions. The chemical reactions that comprise the biochemical network are inherently stochastic in nature. In this thesis, I have addressed two main issues regarding stochastic modeling and simulation of biochemical kinetics. First, the biochemical reactions involved in cellular processes are often multi component and involve complex multi dimensional dynamics. It is often in the interest of computational simplification and even an analytical closed form solution to reduce the dimensionality of the kinetic model. In chapter 2, I have evaluated the precision of a well known method for dimensionality reduction (QSSA) in

case of stochastic simulations. Second, the stochastic simulations of a complex biomolecular network may be extremely dependent of choice of parameters. Many of the model parameters are not available from experiment. So in chapter 3, I have applied the stochastic simulation algorithm [15] to a translation based biomolecular loop implicated in learning and memory formation and I have evaluated the previous results about long term maintenance of synaptic plasticity [20] under a stochastic simulation framework.

Chemical reactions involving small number of reacting molecules are essentially stochastic in nature. The mass action approach to obtain a deterministic approximation works well in estimating the mean concentrations of reacting species over time, when the system is well mixed and the number of molecules is large. In certain reaction systems, where the stochastic fluctuations are large, the distribution around the mean are of interest.

Most protein networks are quite complex, even in their deterministic limit, and their complexity and computational cost increase significantly once stochastic fluctuations are taken into account. For many years a dimensionality reduction technique, the quasi-steady state assumption has been applied to deterministic models, and recently it has been suggested that similar techniques can be useful when applied to stochastic systems as well [17, 16]. Whether this approach yields precise results is the question addressed in chapter 2 of this thesis.

The canonical technique for simulating stochastic chemical reactions is the exact stochastic simulation algorithm, proposed by Gillespie in 1977. This

algorithm can be very computationally expensive. Various approximations have been proposed to speed it up, and one such approach, proposed in 2003 by Rao and Arkin suggested using the QSSA approach, and using the reduced master equation to derive a Gillespie-like algorithm. Rao and Arkin (2003) [17] have shown that this method significantly reduces the computational cost, and at the same time produces a good estimate of the mean of the full model. The errors in predicting the mean with this Gillespie-like algorithm mirror the errors that the QSSA produces in deterministic simulations. However, in that paper they do not examine if this Gillespie-like approach estimates well other properties of the distribution of the chemical species.

We have shown two examples where a QSSA based reduction of the Gillespie scheme leads to significant errors in the distribution of stochastic fluctuations around the mean. The first example we examined is closely related to the simple catalytic reaction scheme, with one additional reaction that avoids a trivial fixed point. For this reaction we show that by applying the QSSA we obtain large errors in the distribution of the substrate. We show this both using an analytical solution of the master equation, and by preforming stochastic simulations of the full and reduced models. Surprisingly we find that the reduced model has much higher variance than the original model. This is the case even though we tested the model in the regime where the QSSA should be valid. We also found no apparent relationship with how well the QSSA approximates the determinisitc dynamics, and how well the QSSA in stochastic simulations approximates the substrate distribution.

Our other example is a bi-stable autophosphorylation loop, inspired by a model proposed by Lisman in 1985, to account for the stability of memory. The model differs from the original model in having an additional constitutive kinase activity to avoid an absorbing state [3] at the low fixed point, and in that we systematically applied the QSSA approximation to reduce this model to a 1-D system. The stochastic simulations of both the full and reduced models show spontaneous transitions between a 'Down' and an 'Up' state. For both the full and reduced model these states have the same value. However, here too the distributions of the reduced model are significantly different than the distributions obtained by simulating the reduced model. Most significantly the residence times in both the 'Down' and 'Up' states in the reduced model are two orders of magnitude shorter than the residence times of the full model. This result implies that such reductions might not be useful if for example the statistic of biological significance is the residence time.

Another approach for dimensionality reduction in chemical reactions is the total quasi-steady state approach [13]. This approach when applied to a simple catalytic reaction preforms a change of variables, and assumes a time scale separation not between the original reactions, but for the transformed set of reactions. For deterministic cases this approach extends the range over which the quasi-steady state assumptions can be applied, at the cost of producing more complex equations that do not have an intuitive closed form. Recently this approach was extended to stochastic systems, by using again a Gillespie-like algorithm directly for the equations produced by the to-

tal QSSA [16]. Applying this approach to various mono-stable and bi-stable systems produced results that seem very close to those produced by the full system. We did not examine here the impact of the total QSSA approximation, and this should be carried out in subsequent work. However, our results are qualitatively both similar and different from the results of using total QSSA approach for dimensionality reduction in stochastic simulations. First, in our simple enzymatic reaction we added an additional step, not included by the Barik et al (2008) paper. Therefore, our system has an additional parameter. Nevertheless, we have simulated the catalytic reaction system with all parameters except for one being identical to those on Barik et. al. (2008) (Figure 2.3 - star symbols). For few of those simulations (i.e. for few values of our parameter $k3$ in Eqn 2.7) we indeed find a relatively good agreement between the full and reduced models, even without applying directly the total QSSA approach. For other choices of $k3$, the agreement between full and reduced stochastic distributions is not good. In our simulations of the bi-stable system we find a big difference in the residence times between the full and reduced models, and indeed this is consistent with results obtained for some sets of parameters for the different bi-stable system modeled by Barik et al (2008) (Figure 10 in Barik et al (2008)). In a sense it is not surprising that the biggest differences between the full and reduced models are obtained for the residence time statistics. Transitions between these two states can be thought of as jumps between two energy minima, triggered by stochastic fluctuations. If the two states are seen as separate in the stochastic simulations, implies that

the tails of the distribution produce these transitions. The mass of such tails can indeed be significantly different even if the parameters of the distribution are only moderately different, resulting in a large difference in the residence time statistics.

Several other approaches have been suggested for accelerating stochastic simulations. One well known approach, the tau-leap was proposed Gillespie and co-workers [25, 26]. Related techniques are various reduction methods such as slow-scale stochastic simulation algorithm, accelerated stochastic simulation algorithm etc. The reader is redirected to Gillespie (2007) [27] and Gillespie et. al. (2009) [28] for a review of these algorithms. Other approaches, are the hybrid simulation algorithms [29]. According to such approaches reactions are partitioned into fast and slow, fast reactions are simulated with deterministic or fast approximate stochastic techniques, whereas slow reactions still use the complete exact stochastic scheme. Such results too can speed up run times, and produce decent approximations for large time scale separations.

The traditional QSSA approach not only speeds up simulations times. It also has the advantage of producing reduced dimensional, closed form and intuitive sets of differential equations. We have given two examples here of the catalytic reaction and another of a simple bistable system based on autophosphorylation. For the catalytic reaction scheme we have a closed form solution of the 1-d reduced deterministic model, but more importantly can also analytically find the solution of the chemical master equation. In the bi-stable system, we have a closed form solution of the deterministic 1-d reduced model,

which intuitively can explain the origin of the bi-stability.

Long term synaptic plasticity is maintained through persistent activity of certain proteins. An example is the activity of αCamKII which phosphorylates the AMPA receptors for trafficking to the cell membrane in the post synaptic neuron. A reaction network consisting of a αCamKII autophosphorylation loop coupled with a CPEB1 mediated up-regulation of αCamKII synthesis via translation that acts as a bistable system is implicated in maintenance of long term plasticty. A model for this reaction was successfully able to explain the seemingly different outcome of application of protein synthesis inhibitors in the induction and maintenance phases of L-LTP [20].

The previous model of the bi-stable CaMKII-CPEB1 feedback loop was based on deterministic chemical kinetics. However, the chemical reactions in this model are taking place in the small volume of the post synaptic spine. In such a small volume, only a small number of molecules are involved in the reaction process. This lends inherent stochasticity to the process. In this thesis, we examined the impact of the stochastic dynamics in order to see if the system can remain essentially bi-stable and if the transition times between states are long enough to account for the stability of memory. In order to implement stochastic dynamics we improved the translation model, and examined when it is appropriate to model it as an elementary reaction, as appropriate for the Gillespie algorithm [15]

In our current model of translation, we have explicitly modeled the mRNA molecule and the polyribosome complex. When the CPEB1 molecule

binds to a mRNA, it gets activated to bind to the polyribosome complex. If the CPEB bound was phosphorylated then the activation is strong and if the CPEB1 bound was dephosphorylated, the activation is weak. The strongly activated mRNA is more likely to bind to the polyribosome complex as compared to the weakly activated mRNA.

The use of Gillespie algorithm requires that each reaction be a Poisson process such that the time interval between successive instances of the reaction are distributed exponentially. Elementary reactions are Poisson in nature. The process of translation as a whole is not an elementary reaction. However, we have shown that if several asynchronous polyribosomes act together, the overall inter-protein production interval that they produce can be closely approximated by an exponential distribution. Thus in this case, the use of Gillespie algorithm for modeling translation was found appropriate. The details of this is provided in the Appendix.

Our deterministic model proposed that the different concentrations of αCamKII in long-term potentiated and depotentiated synapses was due to the bistable system. However, strictly, the concept of bistability exists only for deterministic systems as in a stochastic systems, there are transitions from one state to another due to fluctuations. We tested the effect of level of overall stochasticity in the system on the maintenance of long term plasticity. For a system with sufficiently large fluctuations, the long term plasticity as observed in a deterministic model would be reversed merely because of stochastic fluctuations. In figure 3.3, we ran the simulations for three different reaction

volumes. The overall rate of protein production, which is an important determiner of stochastic fluctuations in translation, is adjusted to that observed in experiment [23]. It can be seen that for a reaction volume that is comparable to the volume of a typical postsynaptic spine [30], the concentration of total αCamKII is maintained at the level required for LTP for at least 30 hours. In other runs of the same simulation (not shown), we could not observe a reversal of the upstate even even after 120 hours.

The level of stochasticity in our simulation depends on the parameters of the model. The rate constants of the biochemical reactions involved in our model are very difficult to be measured in experiment. In figure 3.4, we have shown the results of three stochastic simulations with different values of the rate constant for phosphorylated CPEB1 binding with mRNA. The cumulative distributions for the residence time in the upstate shows that the level of stochasticity which affects the over all equilibrium behavior of the system is very sensitive to the choice of free parameters.

The model in this thesis discusses the importance of stochasticity for studying the maintenance of long term plasticity only in case of translation and post-translation modulation of plasticity. Translation based local regulation of synaptic plasticity seems plausible because electron imaging experiments have found presence of translation machinery in the spine of the post synaptic neuron [30]. Stochasticity at the level of gene-expression may also play a role in over-all availability of mRNA molecules in the spine. This may also affect the overall stability of the maintenance of long term synaptic plasticity.

Appendices

Appendix A

Mathematical methods

A.1 Energy Function

The QSSA based reduction of the deterministic system gives us the 1-d equation:

$$\frac{\mathrm{d}K_P}{\mathrm{d}t} = f(K_P)$$

The energy function for this equation as a function of K_P is given by:

$$U(K_P) = \int_0^{K_P} f(x)\mathrm{d}x \tag{A.1}$$

The energy function in Figure 2.4 is generated by symbolic integration of the right hand side of equation 2.33.

A.2 Normalized Mean Square Error function

Figure 2.3 shows the comparison between the accuracy of the quasi-steady state approximation for deterministic simulations and the reduced chemical master equation for the stochastic simulations. On the x-axis is a measure of the accuracy of quasi-steady state approximation of the deterministic approximation. This measure is the mean-square error between a deterministic simulation implementing quasi-steady state and a deterministic simulation of

the full model during the transient, i.e. before the steady state is reached (The error is zero after the steady state is reached). The mean square error is normalized by the time taken for the transient to reach the steady state. This ensures that the measure is invariant to the rate of convergence the steady state. The error is measured relative to the concentration of the full model. This ensures that the measure is scale-invariant. Mathematically it takes the form:

$$\text{NMSE} = \left\langle \frac{1}{t_N} \sum_{i=1}^{N} \left(\frac{\hat{x}(t_i)_{full} - \hat{x}(t_i)_{reduced}}{\hat{x}(t_i)_{full}} \right)^2 \right\rangle \tag{A.2}$$

where t_N is the time it takes for a particular trial to reach the steady state, $\hat{x}_{full_i}$ and $\hat{x}_{reduced_i}$ are the intrapolated x vectors on an equi-spaced time grid from 0 to t_N for full and reduced deterministic simulations respectively. This error in transients is averaged over hundred trials with randomized initial conditions. A smaller value of NMSE means that the deterministic quasi-steady state transient better approximates the deterministic full model transient for a particular set of parameters.

A.3 Kullback-Leibler divergence

On the y-axis of Figure 2.3 is the Kullback-Leibler Divergence between the substrate distribution at steady state for the complete reaction system and the reduced model. KL Divergence is used as a measure of difference between the stochastic distributions of substrate in the full and reduced model. KL

Divergence is given by the following formula:

$$\text{KL Divergence} = \int_{-\infty}^{\infty} p_{full}(x) \log \frac{p_{full}(x)}{p_{reduced}(x)} \mathrm{d}x \tag{A.3}$$

where $p_{full}(x)$ is the probability density of the substrate distribution for the full model and $p_{reduced}(x)$ is the probability density of the substrate distribution for the reduced model.

A.4 Proof for stationary distribution for the chemical master equation in Eqn 2.10

For the chemical master equation as in Eqn 2.10 with the boundary conditions as described by Eqn 2.11 - 2.12, at steady state,

$$\frac{\mathrm{d}\rho(n)}{\mathrm{d}t} = 0 \qquad \forall \quad n = 0, 1, 2 \ldots \infty \tag{A.4}$$

Proof by induction: For $n = 0$,

$$\begin{aligned} \frac{\mathrm{d}\rho(0)}{\mathrm{d}t} = 0 \quad &\Rightarrow \quad r(1)\rho(1) - r(0)\rho(0) + g(-1)\rho(-1) - g(0)\rho(0) = 0 \\ &\Rightarrow \quad \rho(1) = \frac{g(0)}{r(1)}\rho(0) \end{aligned} \tag{A.5}$$

Let's assume that for some $k \in [0, \infty]$,

$$\rho(k) = \frac{g(k-1)}{r(k)}\rho(k-1) \tag{A.6}$$

For steady state from Eqn A.4 for $n = k$,

$$\frac{\mathrm{d}\rho(k)}{\mathrm{d}t} = 0 \quad \Rightarrow \quad r(k+1)\rho(k+1) - r(k)\rho(k) + g(k-1)\rho(k-1) - g(k)\rho(k) = 0$$

$$\text{Using Eqn A.6,} \quad \Rightarrow \quad \rho(k+1) = \frac{g(k)}{r(k+1)}\rho(k) \tag{A.7}$$

Therefore, by induction,

$$\rho(n) = \frac{g(n-1)}{r(n)}\rho(n-1) \qquad \forall \quad n = 1, 2, 3 \ldots \infty \tag{A.8}$$

This recursive relationship can be used to find all the ρ's for $n = 0, 1, 2 \ldots \infty$ by choosing an arbitrary $\rho(0)$ and readjusting all the ρ's by normalization.

Appendix B

Motivation for using Gillespie Algorithm for modeling translation

The stochastic simulation algorithm developed by Gillespie [15] treats every reaction as a poisson-like process with inter-reaction interval sampled from an exponential distribution. In this paper, we have simulated the process of translation, i.e. binding of the activated mRNA to polyribosomes to give rise to a protein molecule in its native state, as a single step in the Gillespie algorithm. The actual process is quite more complex and involves several steps. In this section, we will provide some motivation as to why we have considered translation as a single step in the Gillespie algorithm. The use of Gillespie algorithm for translation could be justified, if the the inter-protein-production interval in the reaction volume is exponentially distributed as shown in Figures B.1 and B.2. We show here, that is indeed the case, in the presence of several protein production machineries running in desynchronized state with each other.

The process of translation takes place in a polyribosome complex. This complex consists of a activated mRNA molecule on which the ribosome molecules traverse from the 5' end to the 3' end, producing a amino acid chain

that detaches at the end of the mRNA molecule when the ribosome falls off. Thus translation broadly involves three steps. In the first step, a free ribosome molecule binds to the polyribosome complex. This process is essentially poisson like and the inter-reaction interval is exponentially distributed. After a ribosome binds to the mRNA chain, it starts to move up towards the 3' end producing a amino-acid chain. Another ribosome cannot bind to the mRNA 5' end until the first ribosome has already moved up some distance. In the second step, the ribosome molecule traverses the length of the mRNA chain from the 5' end to the 3' end. The time a ribosome molecule takes to complete its journey along the mRNA molecule can be assumed to have a gaussian distribution with mean residence time obtained by multiplying the mean step time with the length of mRNA. The third step is the detachment of the ribosome and amino acid chain from the mRNA molecule. This step is also poisson-like with inter-reaction interval exponentially distributed [31].

Figure B.1 shows a schematic of the time line of protein production at a single polyribosome complex. Using the scheme described below the figure, the inter-protein-production time interval distribution is calculated for one polyribosome complex. The protein release events from several such polyribosome complexes, simulated independently of each other, are interleaved on a single protein release timeline to obtain the inter-protein-production time interval distribution in the presence of more than one polyribosome complex in desynchronized state with each other.

Figure B.2 shows the inter-protein-production time interval distribution

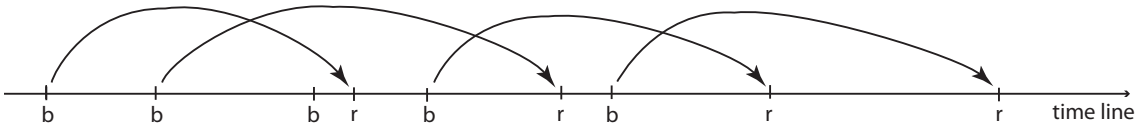

Figure B.1: Schematic of the time line of protein production at a single polyribosome complex. 'b' on the timeline indicates ribosome binding events. Consecutive binding events are separated by inter-binding time intervals that are generated from an exponential distribution whose mean depends on the number of free ribosome molecules. Every binding event is used as a root for generating a protein release event, 'r'. The arrow traces a particular ribosome from its binding event to its release event. The time interval is the sum of a time interval sampled from a gaussian distribution (for elongation step) and a exponential distribution (for termination step). The mean elongation time depends on the length of mRNA chain and the mean termination time is fixed. Note that the time interval between successive protein release events form the inter-protein-production time interval distribution

for a single (N = 1) and three desynchronized polyribosome complexes (N = 3). The mean of the exponential distribution for binding is 10 seconds. The refractory time is also 10 seconds. The mean and standard deviation of the Gaussian distribution for elongation times are 75 seconds and 8.66 seconds respectively. The mean of exponential distribution for termination time is 0.1 seconds. These means are consistent with [23]. The mean of the inter-protein-production time interval for a single polyribosome is 20 seconds and that for three polyribosomes is 6.67 seconds. For three polyribosome complexes, the inter-protein-production interval is quite well approximated by an exponential distribution. Thus the use of Gillespie algorithm to simulate the protein-production step is somewhat justified.

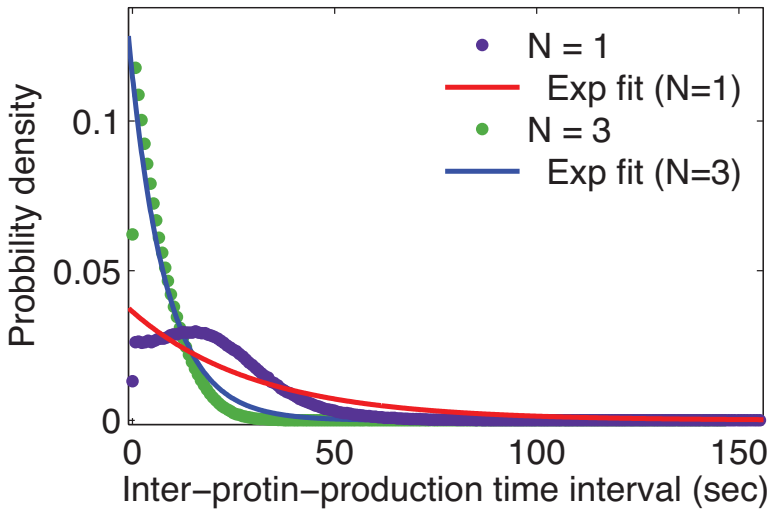

Figure B.2: Comparison of the inter-protein-production time interval distribution for a single (N = 1) and three desynchronized polyribosome complexes (N = 3). An exponential has been fit for both N=1 and N=3. The exponential fit for N=3 is much better (R^2 value = 0.95; $f(t) = 0.1149\mathrm{e}^{-0.1109t}$) as compared to that for N=1 (R^2 value = 0.8584; $f(t) = 0.03631\mathrm{e}^{-0.03251t}$)

Appendix C

Model reactions and parameter table

The reaction scheme outlined in figure 3.1 represents the following set of chemical reactions:

$$
\begin{aligned}
X + U &\underset{k_{1010}}{\overset{k_{10}}{\rightleftarrows}} X_P \\
X + X_P &\underset{k_2}{\overset{k_1}{\rightleftarrows}} C_1 \overset{k_3}{\longrightarrow} 2X_P \\
X_P + P &\overset{k_4}{\longrightarrow} X \\
Y + X_P &\underset{k_{55}}{\overset{k_5}{\rightleftarrows}} C_2 \overset{k_6}{\longrightarrow} Y_P + X_P \\
Y_P + P &\overset{k_7}{\longrightarrow} Y \\
M + Y_P &\underset{k_{1111}}{\overset{k_{11}}{\rightleftarrows}} M_A \\
M + Y &\underset{k_{1212}}{\overset{k_{12}}{\rightleftarrows}} M_I \\
M_A + T &\underset{k_{revA}}{\overset{k_{syn1A}}{\rightleftarrows}} C_3 \overset{k_{syn2A}}{\longrightarrow} T + X + M_A \\
M_I + T &\underset{k_{revI}}{\overset{k_{syn1I}}{\rightleftarrows}} C_4 \overset{k_{syn2I}}{\longrightarrow} T + X + M_I \\
X &\underset{\lambda_1}{\overset{\lambda_1}{\rightleftarrows}} X_{Basal} \\
X_P &\underset{\lambda_1}{\overset{\lambda_1}{\rightleftarrows}} X_{P\ Basal}
\end{aligned}
$$

where X is the kinase in dephosphorylated state, X_P is the kinase in phosphorylated state, U is a second-messenger based signal molecule, P is a generic

phosphatase, Y and Y_P are dephosphorylated and phosphorylated forms of an mRNA activator molecule, M, M_A and M_I are the native, strongly activated and mildly activated forms of kinase mRNA and T is free ribosome molecule. The last two reactions represent both protein turnover and diffusion of protein from and into the spine.

The parameters for the above model are as follows:

Parameter	Value	Description
k_{10}	0.005	CamKII binding to signal
k_{1010}	1.0	CamKII unbinding from signal
k_1	0.81	Bimolecular autophosphorylation
k_2	0.002	Dissociation of autophosphorylation complex
k_3	0.83	Autophosphorylation product formation
k_4	0.002	Dephosphorylation of CamKII
k_5	0.2	Bimolecular binding of CPEB1 to CamKII
k_{55}	0.01	Dissociation of CPEB1-CamKII complex
k_6	0.21	Phosphorylation of CPEB1
k_7	0.02	Dephosphorylation of CPEB1
k_{11}	0.21	Binding of active CPEB1 to mRNA
k_{1111}	1	Unbinding of active CPEB1 from mRNA
k_{12}	0.0001	Binding of inactive CPEB1 to mRNA
k_{1212}	1	Unbinding of inactive CPEB1 from mRNA
k_{syn1A}	0.1	Active mRNA binding with ribosome complex
k_{revA}	0.001	Unbinding of active mRNA from ribosome complex
k_{syn2A}	0.4	Synthesis of new CamKII from active mRNA
k_{syn1I}	0.005	Inactive mRNA binding with ribosome complex
k_{revI}	0.001	Unbinding of inactive mRNA from ribosome complex
k_{syn2I}	0.0021	Synthesis of new CamKII from inactive mRNA
λ_1	0.00099	Turnover for dephosphorylated CamKII
λ_2	0.00099	Turnover for phosphorylated CamKII
U	0.00005	Fixed concentration of signal molecule
P	0.0148	Fixed phosphatase concentration
M	0.0015	Fixed mRNA concentration
X_{Basal}	0.01	Basal dephosphorylated CamKII
$X_{P\ Basal}$	0.01	Basal phosphorylated CamKII
X_{Total}	10	Total concentration of Kinase at t_0
Y_{Total}	28.53	Total concentration of mRNA activator molecule

Bibliography

[1] Natal A. W. van Riel. Dynamic modeling and analysis of biochemical networks: mechanism based models and model-based experiments. *Briefings in bioinformatics*, 7(4):364–374, 2006.

[2] L. A. Segel. On the validity of steady state assumption of enzyme kinetics. *Bull. Math. Biol.*, 6:579–593, 1988.

[3] J. E. Lisman. A mechanism of memory storage insensitive to molecular turnover: A bistable autophosphorylating kinase. *Proc. Natl. Acad. Sci. USA*, 82:3055–3057, 1985.

[4] L. Cai, N. Friedman, and X. S. Xie. Stochastic protein expression in individual cells at the single molecule level. *Nature*, 440:358–362, 2006.

[5] L. C. Yeung, Castellani G. C., and Shouval H. Z. Analysis of the intraspinal calcium dynamics and its implications for the plasticity of spiking neurons. *Phys Rev E.*, 69:011907, 2004.

[6] H. Z. Shouval and Kalantzis G. Stochastic properties of synaptic transmission affect the shape of spike time-dependent plasticity curves. *J Neurophysiol.*, 93:1069–73, 2005.

[7] Y. Cai, J. P. Gavornik, L. N. Cooper, L. C. Yeung, and Shouval H. Z. Effect of stochastic synaptic and dendritic dynamics on synaptic plasticity in visual cortex and hippocampus. *J Neurophysiol.*, 97:375–86, 2007.

[8] Keller. D. X., K. M. Franks, T. M. Jr Bartol, and T. J. Sejnowski. Calmodulin activation by calcium transients in the postsynaptic density of dendritic spines. *PLoS One*, 3:e2045, 2008.

[9] S. Zeng and W. R. Holmes. The effect of noise on CaMKII activation in a dendritic spine during LTP induction. *J Neurophysiol.*, 103:1798–808, 2010.

[10] Y Kubota and M. N. Waxham. Lobe specific Ca2+-calmodulin nanodomain in neuronal spines: a single molecule level analysis. *PLoS Comput Biol.*, 6:e1000987, 2010.

[11] P. Miller, A. M. Zhabotinsky, J. E. Lisman, and Wang X. J. The stability of a stochastic camkii switch: dependence on the number of enzyme molecules and protein turnover. *PLoS Biol.*, 3:e107, 2005.

[12] M. Graupner and N. Brunel. Stdp in a bistable synapse model based on camkii and associated signaling pathways. *PLoS Comput Biol.*, 3:e221, 2007.

[13] J. A. Borghans, R. J. de Boer, and L. A. Segel. Extending the quasi-steady state approximation by changing variables. *Bull. Math biol.*, 58:43–63, 1996.

[14] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry, Third Edition*. North Holland Public Library, 2007.

[15] D. T. Gillespie. Exact stochastic simulations of coupled chemical reactions. *J. Phys. chem.*, 115:1716–1733, 1977.

[16] D. Barik, M. R. Paul, W. T. Baumann, Y. Cao, and J. T. Tyson. Stochastic simulations of enzyme-catalyzed reactions with disparate timescales. *Biophysical Journal*, 95:3563–3574, 2008.

[17] C. V. Rao and A. P. Arkin. Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the gillespie algorithm. *J. Chem. Phys.*, 118:4999–5011, 2003.

[18] T.V. Bliss and G.L. Collingridge. A synaptic model of memory: longterm potentiation in the hippocampus. *Nature*, 361:31–39, 1993.

[19] T.P. Feng. The involvement of pkc and multifunctional cam kinase ii of the postsynaptic neuron in induction and maintenance of long-term potentiation. *Prog Brain Res*, 105:55–63, 1995.

[20] N. Aslam, Y. Kubota, D. Wells, and H.Z. Shouval. Translational switch for long-term maintenance of synaptic plasticity. *Molecular Systems Biology*, 5:284, 2009.

[21] N. Otmakhov, L.C. Griffith, and J.E. Lisman. Postsynaptic inhibitors of calcium/calmodulin-dependent protein kinase type ii block induction but

not maintenance of pairing-induced long-term potentiation. *J Neurosci*, 17:5357–5365, 1997.

[22] M. Sanhueza, C.C. McIntyre, and J.E. Lisman. Reversal of synaptic memory by ca2+/calmodulin-dependent protein kinase ii inhibitor. *J Neurosci*, 27:5190–5199, 2007.

[23] H. Zouridis and V. Hatzimanikatis. A model for protein translation: polysome self-organization leads to maximum protein synthesis rates. *Biophys J*, 92(3):717–730, 2007.

[24] P. Muller. Glossary of terms used in physical organic chemistry (iupac recommendations 1994). *Pure Appl. Chem.*, 66(5):1077–1184, 1994.

[25] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chem. Phys.*, 115:1716–33, 2001.

[26] Y. Cao, D. T. Gillespie, and L. R. Petzold. Efficient stepsize selection for the tau-leaping simulation method. *J. Chem. Phys.*, 124:044109, 2006.

[27] D. T. Gillespie. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58:35–55, 2007.

[28] D. T. Gillespie, Y Cao, K. R. Sanft, and L. R. Petzold. The subtle business of model reduction for stochastic chemical kinetics. *J. Chem. Phys.*, 130:064103, 2009.

[29] H. Salis and Y. Kaznessis. Equation-free probabilistic steady-state approximation: dynamic application to the stochastic simulation of biochemical reaction networks. *J. Chem. Phys.*, 123:214106, 2005.

[30] L.E. Ostroff, J.C. Fiala, B. Allwardt, and K.M. Harris. Polyribosomes redistribute from dendritic shafts into spines with enlarged synapses during ltp in developing rat hippocampal slice. *Neuron*, 35:535–545, 2002.

[31] Ashok Garai, Debanjan Chowdhury, Debashish Chowdhury, and T.V. Ramakrishnan. Stochastic kinetics of ribosomes: single motor properties and collective behavior. *arXiv*, 0903.2608v3, 2009.

Vita

Animesh Agarwal was born in India in 1986. After completion of his high school education from Delhi Public School, Bokaro he received the Bachelor of Technology and Master of technology degrees from Indian Institute of Technology Madras, Chennai in July 2009. During his undergraduate studies, he worked on several aspects of Bioengineering and Chemical engineering (reaction kinetics and process engineering) at prestigious research institutions including Indian Institute of Science (IISc), Bangalore, India, Max Planck institute, Germany and Centro Nacional de Biotecnologia (CNB-CSIC), Madrid, Spain. With a penchant for further research, in August 2009, he joined the master's degree program in the Biomedical Engineering department at The University of Texas At Austin. During his master's he was involved in a stochastic chemical reaction kinetics projects under the supervision of Dr. Harel Shouval. He is an avid reader and an adventure travel enthusiast.

Contact: animesh.agarwal@utexas.edu

This thesis was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.